

## Supplementary Materials for

### Trans-ethnic genome-wide association study of severe COVID-19

Peng Wu<sup>1,2,\*</sup>, Lin Ding<sup>3,4,\*</sup>, Xiaodong Li<sup>5,6,\*</sup>, Siyang Liu<sup>7,\*</sup>, Fanjun Cheng<sup>8,\*</sup>, Qing He<sup>9,\*</sup>, Mingzhong Xiao<sup>5,6</sup>, Ping Wu<sup>1,2</sup>, Hongyan Hou<sup>2,10</sup>, Minghui Jiang<sup>3,4</sup>, Pinpin Long<sup>4,11</sup>, Hao Wang<sup>4,11</sup>, Linlin Liu<sup>12</sup>, Minghan Qu<sup>3,4</sup>, Xian Shi<sup>3,4</sup>, Qin Jiang<sup>4,11</sup>, Tingting Mo<sup>4,11</sup>, Wencheng Ding<sup>1,2</sup>, Yu Fu<sup>1,2</sup>, Shi Han<sup>12</sup>, Xixiang Huo<sup>12</sup>, Yingchun Zeng<sup>12</sup>, Yana Zhou<sup>5,6</sup>, Qing Zhang<sup>5,6</sup>, Jia Ke<sup>5,6</sup>, Xi Xu<sup>5,6</sup>, Wei Ni<sup>5,6</sup>, Zuoyu Shao<sup>5,6</sup>, Jingzhi Wang<sup>5,6</sup>, Panhong Liu<sup>13</sup>, Zilong Li<sup>13</sup>, Yan Jin<sup>14</sup>, Fang Zheng<sup>15</sup>, Fang Wang<sup>9</sup>, Lei Liu<sup>9</sup>, Wending Li<sup>4,11</sup>, Kang Liu<sup>4,11</sup>, Rong Peng<sup>4,11</sup>, Xuedan Xu<sup>4,11</sup>, Yuhui Lin<sup>4,11</sup>, Hui Gao<sup>4,11</sup>, Limei Shi<sup>4,11</sup>, Ziyue Geng<sup>4,11</sup>, Xuanwen Mu<sup>4,11</sup>, Yu Yan<sup>3,4</sup>, Kai Wang<sup>3,4</sup>, Degang Wu<sup>3,4</sup>, Xingjie Hao<sup>3,4</sup>, Shanshan Cheng<sup>3,4</sup>, Gaokun Qiu<sup>4,11</sup>, Huan Guo<sup>4,11</sup>, Kezhen Li<sup>1,2</sup>, Gang Chen<sup>1,2</sup>, Ziyong Sun<sup>2,10</sup>, Xihong Lin<sup>16,17,18</sup>, Xin Jin<sup>19,#</sup>, Feng Wang<sup>2,10,#</sup>, Chaoyang Sun<sup>1,2,#</sup>, Chaolong Wang<sup>2,3,4,#</sup>

**Supplementary Table 1. Comparison of top association loci with and without adjustment for age and sex in the analyses of Chinese samples.**

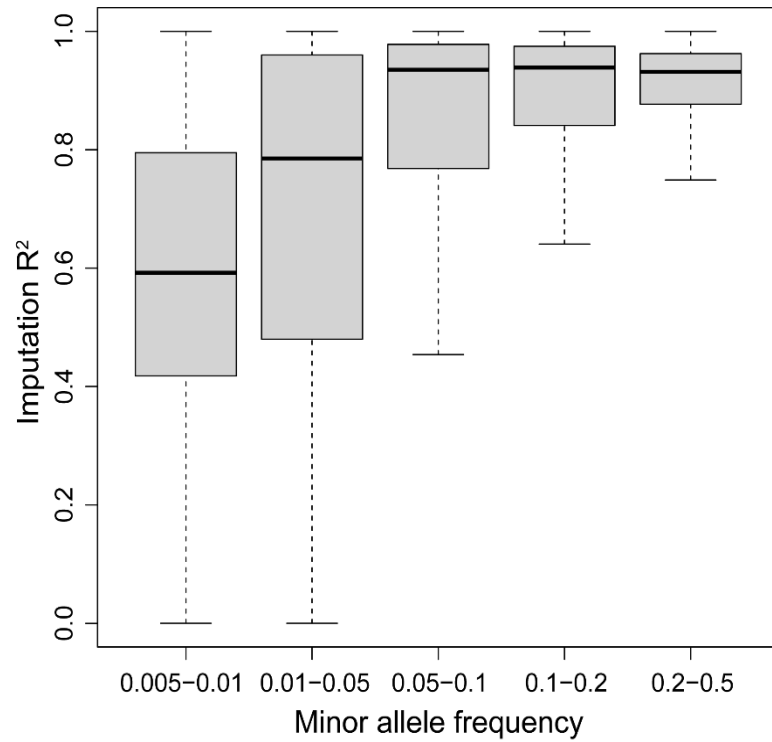
Lead variant	Dataset	No adjustment for age and sex			Adjusting for age and sex		
		OR (95% CI) <sup>†</sup>	<i>P</i>	Heterogeneity	OR (95% CI) <sup>†</sup>	<i>P</i>	Heterogeneity
rs1853837 at 6p21.1	Chinese (GWAS)	1.30 (1.13-1.50)	3.24×10 <sup>-4</sup>		1.33 (1.12-1.57)	9.48×10 <sup>-4</sup>	
	HGI (B2_release3)	1.28 (1.15-1.42)	5.24×10 <sup>-6</sup>		1.28 (1.15-1.42)	5.24×10 <sup>-6</sup>	
	Chinese (WGS)	1.27 (1.07-1.51)	7.06×10 <sup>-3</sup>	<i>I</i> <sup>2</sup> =0.00%	1.30 (1.08-1.57)	6.24×10 <sup>-3</sup>	<i>I</i> <sup>2</sup> =0.00%
	Meta-analysis	1.28 (1.19-1.39)	2.51×10 <sup>-10</sup>	<i>P</i> <sub>het</sub> =0.97	1.29 (1.19-1.40)	4.20×10 <sup>-10</sup>	<i>P</i> <sub>het</sub> =0.93
rs8176719 at 9q34.2	Chinese (GWAS)	1.28 (1.12-1.46)	3.19×10 <sup>-4</sup>		1.15 (0.99-1.35)	6.80×10 <sup>-2</sup>	
	HGI (B2_release3)	1.17 (1.09-1.26)	1.27×10 <sup>-5</sup>		1.17 (1.09-1.26)	1.27×10 <sup>-5</sup>	
	Chinese (WGS)	1.17 (0.98-1.38)	8.03×10 <sup>-2</sup>	<i>I</i> <sup>2</sup> =0.00%	1.17 (0.97-1.41)	1.08×10 <sup>-1</sup>	<i>I</i> <sup>2</sup> =0.00%
	Meta-analysis	1.19 (1.12-1.26)	8.98×10 <sup>-9</sup>	<i>P</i> <sub>het</sub> =0.51	1.17 (1.10-1.24)	5.88×10 <sup>-7</sup>	<i>P</i> <sub>het</sub> =0.99
rs74490654 at 19q13.11	Chinese (WGS)	8.73 (4.14-18.41)	1.22×10 <sup>-8</sup>	-	10.93 (4.56-26.19)	8.22×10 <sup>-8</sup>	-

Notes: †Odds ratio (OR) and 95% confidence interval (CI) of the alternative allele. Meta-analysis is based on the Han-Eskin random-effect method. All association analyses of Chinese samples have adjusted for the top two PCs.

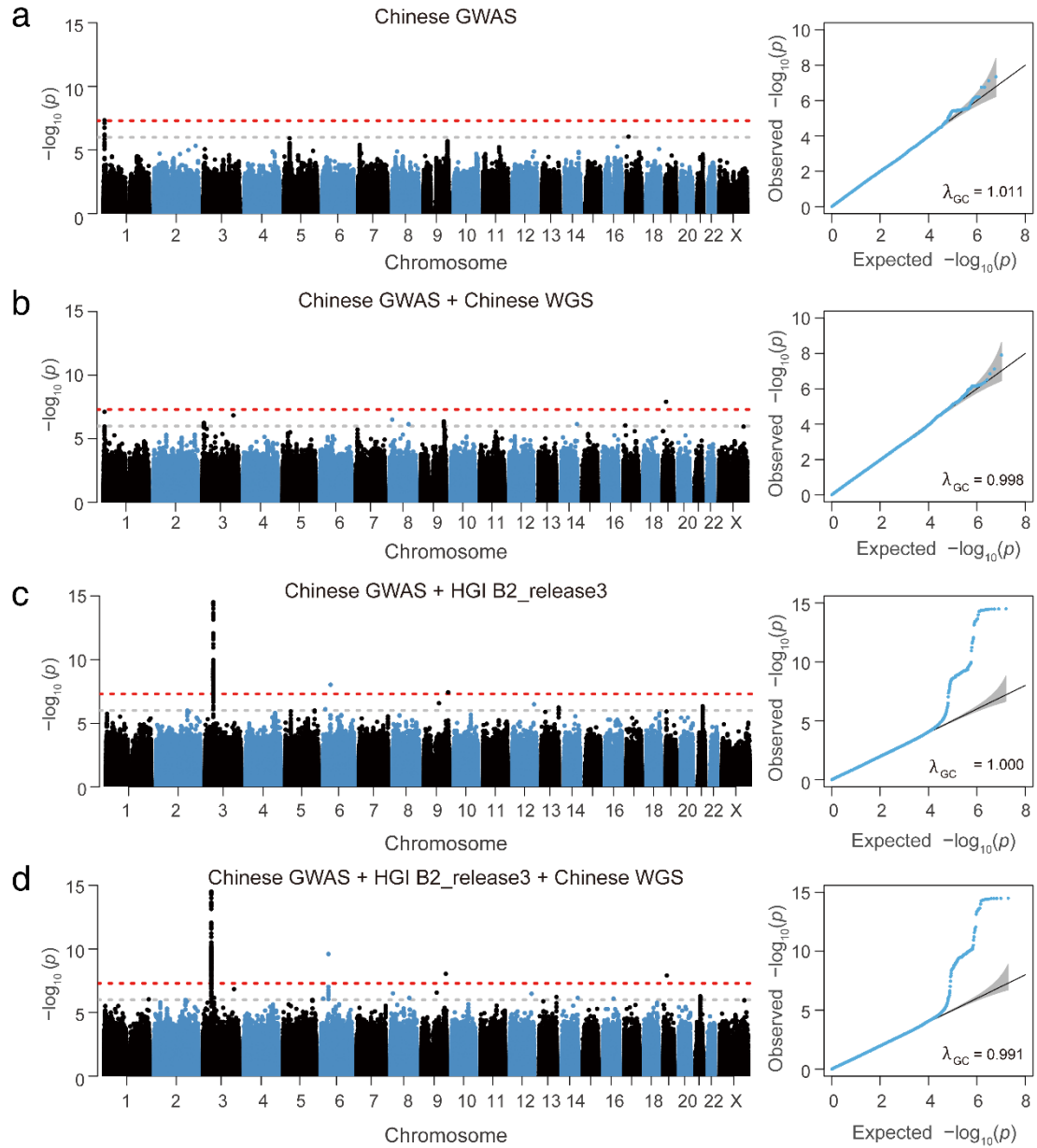
**Supplementary Table 2. Suggestive loci associated with COVID-19 severity ( $P < 10^{-6}$ ).**

Locus	Dataset	Sample size	Lead variant	AF <sup>\$</sup>	OR (95% CI) <sup>†</sup>	<i>P</i>	Heterogeneity
21q22.11 <i>IFNAR2</i>	Chinese (GWAS)	598/2,260	rs1051393	0.610	1.12 (0.97-1.28)	$1.16 \times 10^{-1}$	
	HGI (B2_release3)	3,199/897,488	chr21: 33241950	0.333	1.19 (1.11-1.27)	$1.12 \times 10^{-6}$	
	Chinese (WGS)	474/1,615	T/G	-	-	-	$I^2=0.00\%$
	Meta-analysis	3,797/899,748	Missense		1.17 (1.10-1.25)	$4.33 \times 10^{-7}$	$P_{\text{het}}=0.43$
3p14.2 <i>PTPRG</i>	Chinese (GWAS)	598/2,260	rs672699	0.478	1.04 (0.91-1.19)	$5.80 \times 10^{-1}$	
	HGI (B2_release3)	3,199/897,488	chr3:61768231	0.789	1.19 (1.10-1.29)	$1.36 \times 10^{-5}$	
	Chinese (WGS)	474/1,615	T/A	0.484	1.37 (1.14-1.63)	$5.49 \times 10^{-4}$	$I^2=67.21\%$
	Meta-analysis	4,271/901,363	Intronic		1.18 (1.04-1.34)	$5.58 \times 10^{-7}$	$P_{\text{het}}=0.05$
16q21 <i>ADGRG1</i>	Chinese (GWAS)	598/2,260	rs7499679	0.250	0.85 (0.72-0.99)	$3.80 \times 10^{-2}$	
	HGI (B2_release3)	3,199/897,488	chr16:57636629	0.227	0.86 (0.79-0.92)	$5.92 \times 10^{-5}$	
	Chinese (WGS)	474/1,615	G/A	0.276	0.80 (0.66-0.97)	$2.65 \times 10^{-2}$	$I^2=0.00\%$
	Meta-analysis	4,271/901,363	Intronic		0.85 (0.79-0.90)	$8.09 \times 10^{-7}$	$P_{\text{het}}=0.82$
1q44 <i>HNRNPU</i>	Chinese (GWAS)	598/2,260	rs12130553	-	-	-	
	HGI (B2_release3)	3,199/897,488	chr1:244873270	0.437	1.19 (1.11-1.28)	$4.19 \times 10^{-6}$	
	Chinese (WGS)	474/1,615	T/C	0.338	1.18 (0.99-1.41)	$6.91 \times 10^{-2}$	$I^2=0.00\%$
	Meta-analysis	3,673/899,103	Intergenic		1.19 (1.11-1.27)	$9.17 \times 10^{-7}$	$P_{\text{het}}=0.93$

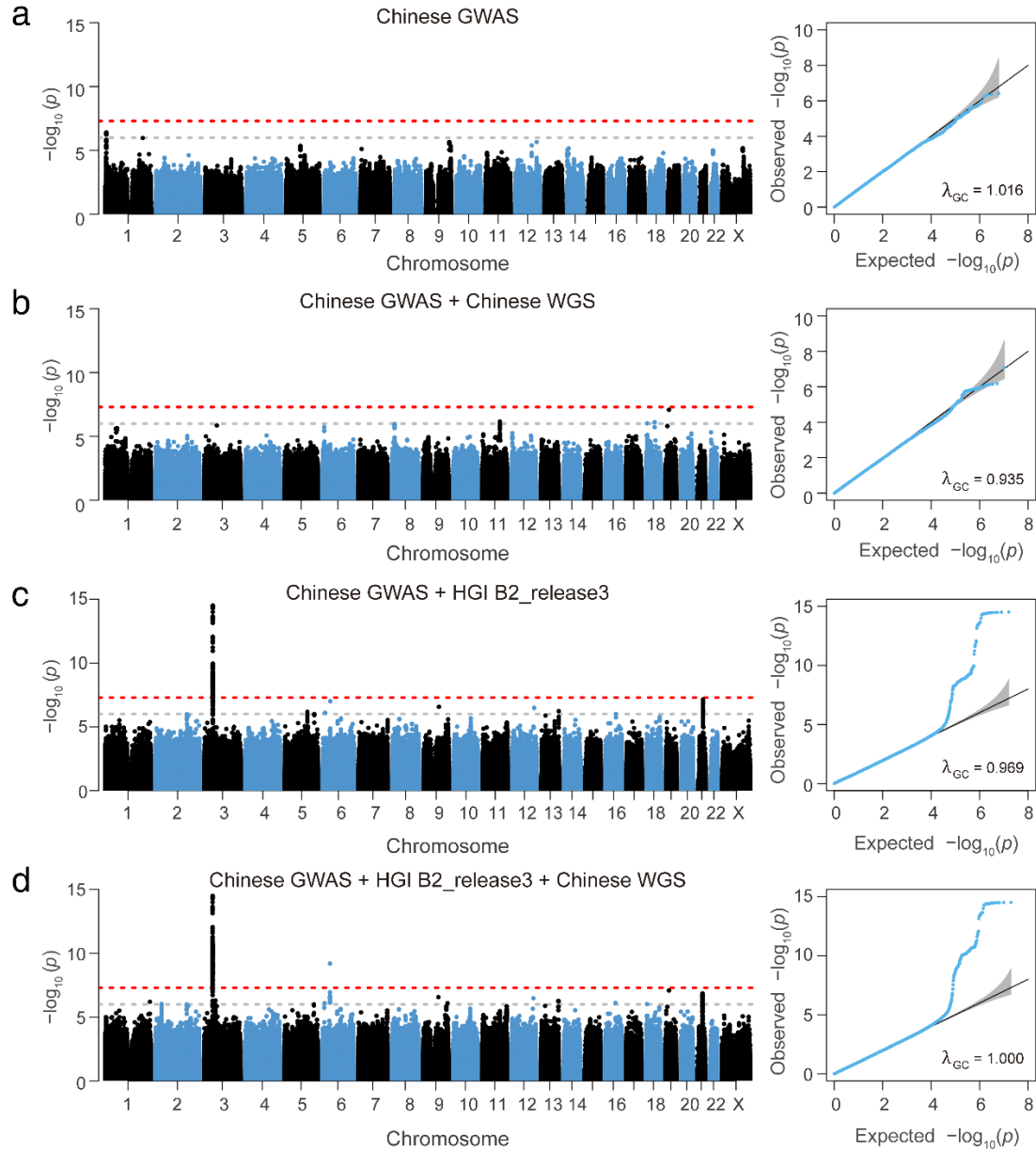
Notes: This table presents loci that are included in at least two datasets and have meta-analysis *P* value between  $5 \times 10^{-8}$  and  $10^{-6}$ . Sample size is presented as number of cases / number of controls. \* Variant with the smallest p value within each locus: rs number, GRCh38 genomic position, reference/alternative alleles, annotation. \$ AF: frequency of the alternative allele: from top to down is the AF in Chinese GWAS controls, the AF in 1KGP European samples, and the AF in Chinese WGS controls. †Odds ratio (OR) and 95% confidence interval (CI) of the alternative allele. Meta-analysis is based on random-effect model.



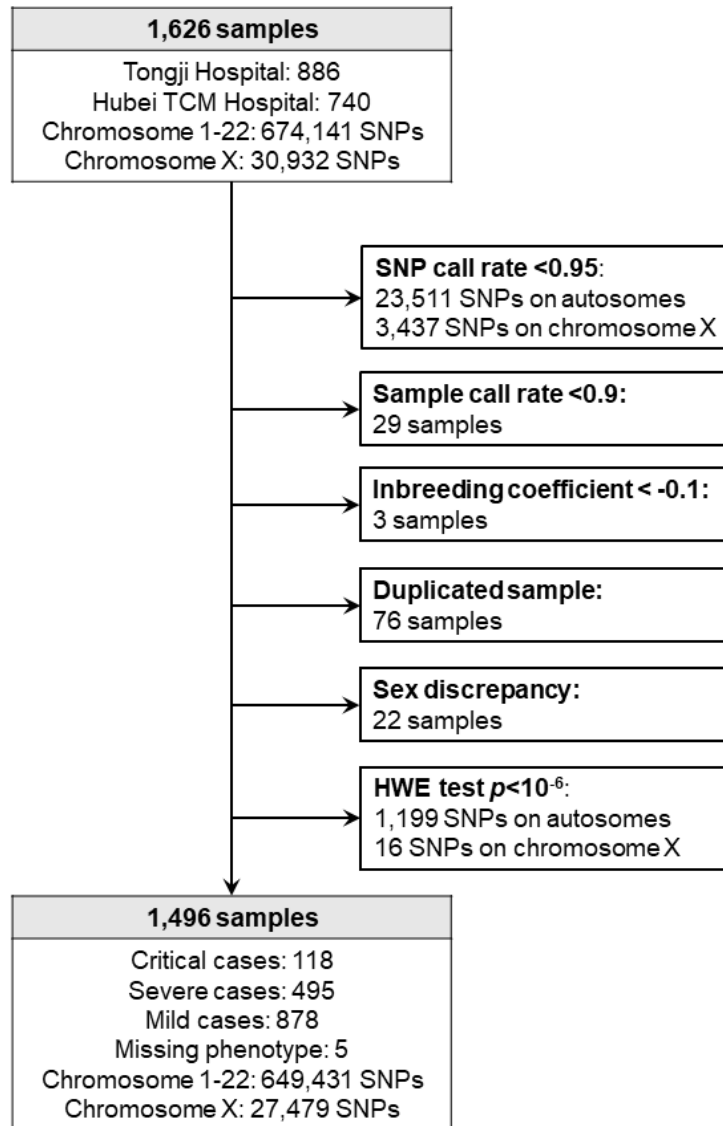
**Supplementary Figure 1. Imputation quality as a function of MAF.** Each box summarizes the imputation  $R^2$  for autosomal SNPs within a MAF bin. The dark horizontal line represents median value, and the grey box represents interquartile range (IQR). Outliers below the lower whiskers ( $1.5 \times \text{IQR}$  below the 25<sup>th</sup> percentile) of the last three bins are not shown.



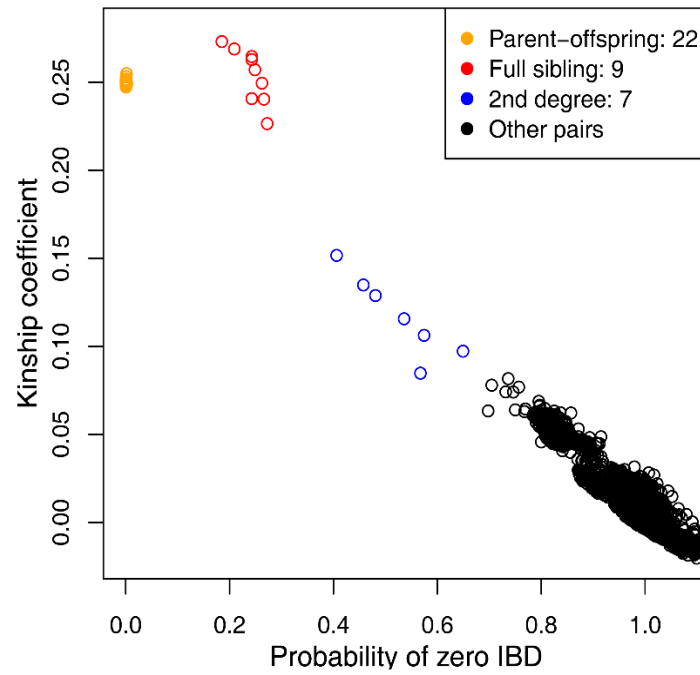
**Supplementary Figure 2. GWAS results for severe COVID-19 with first two PCs included as covariates.** (a) Manhattan and QQ plots for Chinese GWAS. (b) Manhattan and QQ plots for the meta-analysis of Chinese GWAS and Chinese WGS results. (c) Manhattan and QQ plots for the meta-analysis of Chinese GWAS and HGI B2\_release3 results. (d) Manhattan and QQ plots for the meta-analysis of Chinese GWAS, HGI B2\_release3 results, and Chinese WGS. In Manhattan plots, the red dash line indicates genome-wide significance level of  $P=5 \times 10^{-8}$  and the grey dash line indicates suggestive significance level of  $P=10^{-6}$ . In QQ plots, the grey region represents the 95% CI of  $P$  values under the null hypothesis of no association.



**Supplementary Figure 3. GWAS results for severe COVID-19 with first two PCs, age and sex included as covariates.** (a) Manhattan and QQ plots for Chinese GWAS. (b) Manhattan and QQ plots for the meta-analysis of Chinese GWAS and Chinese WGS results. (c) Manhattan and QQ plots for the meta-analysis of Chinese GWAS and HGI B2\_release3 results. (d) Manhattan and QQ plots for the meta-analysis of Chinese GWAS, HGI B2\_release3 results, and Chinese WGS. In Manhattan plots, the red dash line indicates genome-wide significance level of  $P=5 \times 10^{-8}$  and the grey dash line indicates suggestive significance level of  $P=10^{-6}$ . In QQ plots, the grey region represents the 95% CI of  $P$  values under the null hypothesis of no association.

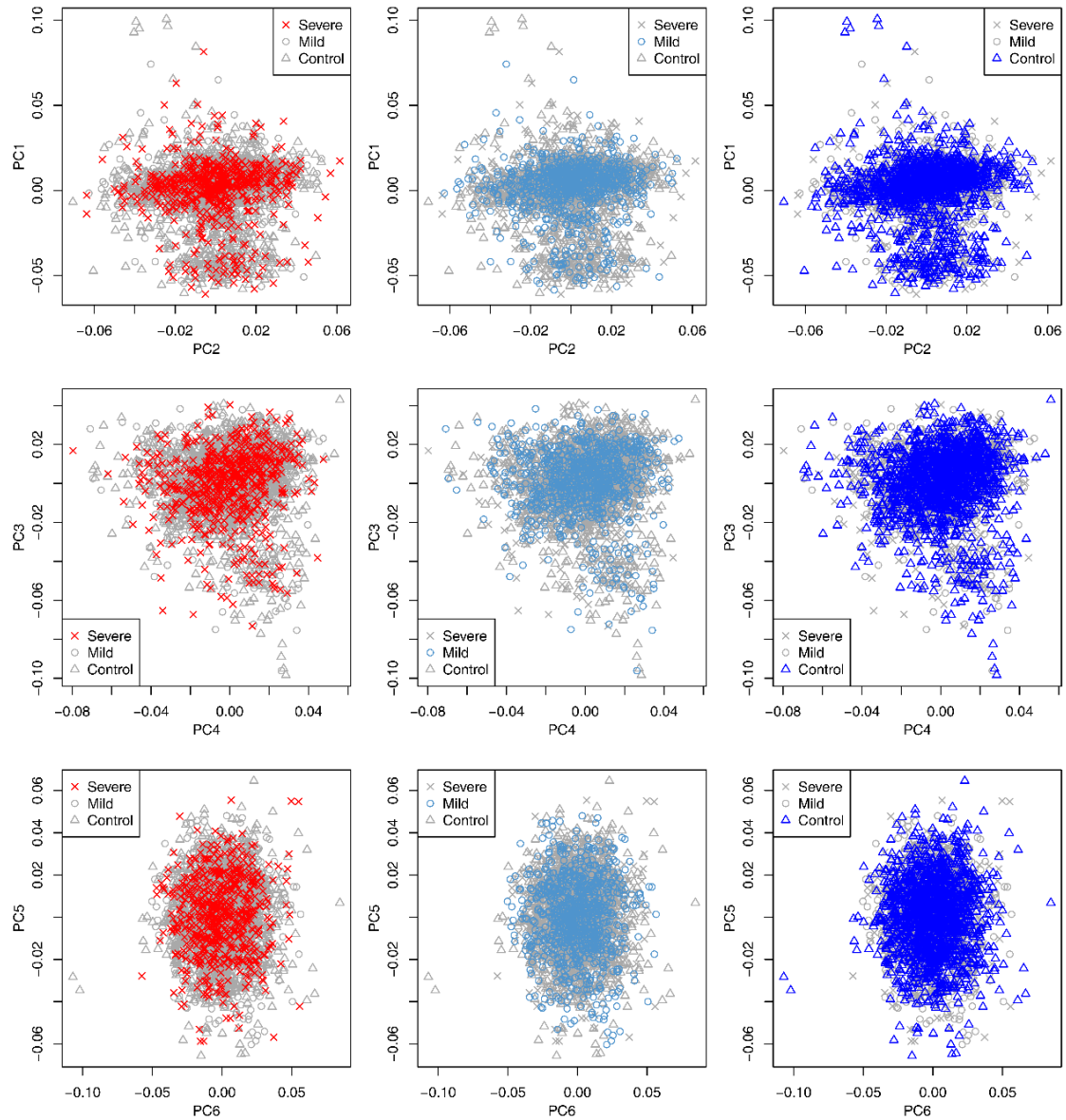


**Supplementary Figure 4. Quality control of the GWAS array data of COVID-19 patients.** Hubei TCM Hospital: Hubei Hospital of Traditional Chinese Medicine.

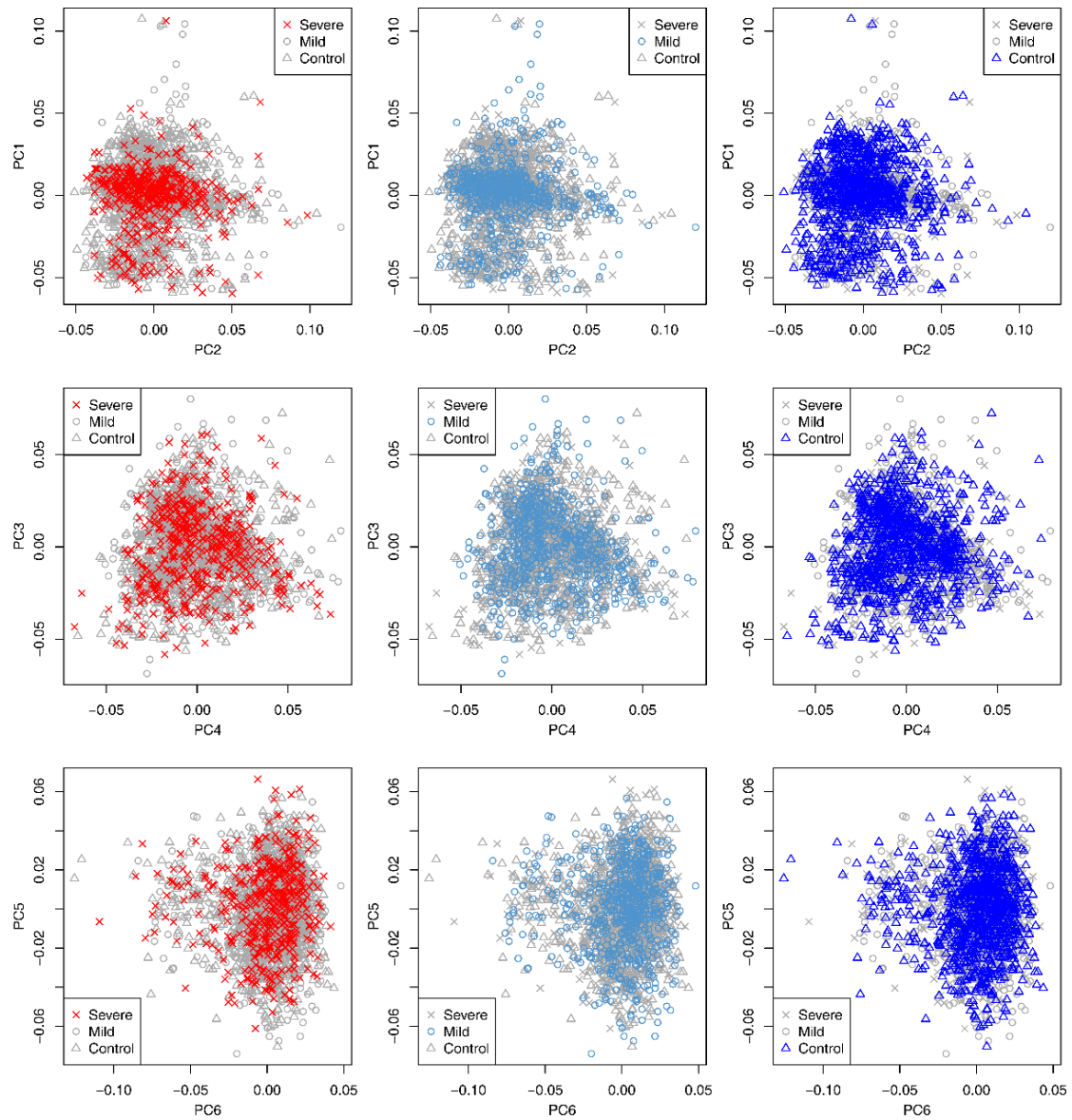


**Supplementary Figure 5. Cryptic relatedness in the GSA genotyped COVID-19 samples.** Numbers of pairs for each relatedness type were presented in the legend. The x-axis is the probability of sharing 0 alleles identical-by-descent (IBD) at a SNP between two individuals.





**Supplementary Figure 6. PCA of the Chinese GWAS samples.** We highlighted samples severe COVID-19, mild COVID-19, and ancestry-matched population controls in columns from left to right, respectively.



**Supplementary Figure 7. PCA of Chinese WGS samples.** We highlighted samples severe COVID-19, mild COVID-19, and ancestry-matched population controls in columns from left to right, respectively.